

CONTROL PROCESSING SYSTEM FOR EXECUTION MULTIPLICITY OF PROCES

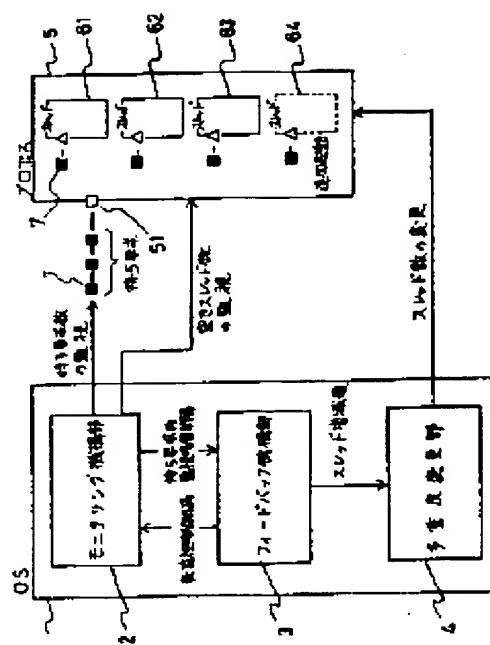
Patent number: JP3040034
Publication date: 1991-02-20
Inventor: SAKAI MASANORI
Applicant: FUJITSU LTD
Classification:
 - international: **G06F9/46; G06F9/46;** (IPC1-7): G06F9/46
 - european:
Application number: JP19890174920 19890706
Priority number(s): JP19890174920 19890706

Report a data error here

Abstract of JP3040034

PURPOSE:To dynamically change the execution multiplicity of a process by changing dynamically the number of threads and a monitoring time interval in accordance with the variation of the number of wait requests and the number of free threads.

CONSTITUTION:An operating system 1 is provided with a monitoring mechanism part 2, a feedback mechanism part 3 and a multiplicity changing part 4. As a result of monitoring by the monitoring mechanism part 2, when the number of wait requests increases (or it is large), the feedback mechanism part 3 derives these values so that the number of threads is increased and a monitoring time interval is shortened. Subsequently, the multiplicity changing part 4 enlarges the execution multiplicity by increasing the number of threads, and also, the monitoring mechanism part 2 monitors a process 5 at a shorter monitoring time interval. The execution multiplicity of the process can be changed dynamically in accordance with a variation of the number of wait requests and the number of free threads.



Data supplied from the esp@cenet database - Worldwide

⑩ 日本国特許庁(JP) ⑪ 特許出願公開
⑫ 公開特許公報(A) 平3-40034

⑬ Int. Cl.⁵
G 06 F 9/46

識別記号 庁内整理番号
3 4 0 A 8945-5B

⑭ 公開 平成3年(1991)2月20日

審査請求 未請求 請求項の数 1 (全7頁)

⑮ 発明の名称 プロセスの実行多重度制御処理方式

⑯ 特 願 平1-174920

⑰ 出 願 平1(1989)7月6日

⑱ 発 明 者 坂 井 正 徳 神奈川県川崎市中原区上小田中1015番地 富士通株式会社
内

⑲ 出 願 人 富士通株式会社 神奈川県川崎市中原区上小田中1015番地

⑳ 代 理 人 弁理士 森 田 寛 外2名

明 細 書

1. 発明の名称

プロセスの実行多重度制御処理方式

2. 特許請求の範囲

要求(7)を多重に実行可能とされ、その実行多重度の数だけのスレッド(61ないし64)を持ち、前記実行多重度を変更可能とされたプロセス(5)を備えたデータ処理システムにおいて、

前記プロセス(5)についての待ち要求の数と空きスレッドの数とをフィードバック機構部(3)において求められた監視時間間隔で監視するモニタリング機構部(2)と、

前記監視の結果に基づいて前記スレッド(61ないし64)の増減数と新たな監視時間間隔とを求めるフィードバック機構部(3)と、

前記スレッド(61ないし64)の増減数に従って前記スレッド(61ないし64)の数を変更することによって、前記実行多重度を変更する多

重度変更部(4)とを設け、

前記待ち要求の数の増加又は減少若しくは空きスレッドの数に応じて、前記多重度変更部(4)が前記スレッド(61ないし64)の数を増加又は減少させることにより前記実行多重度を大きく又は小さくすると共に、前記モニタリング機構部(2)がより短い又はより長い監視時間間隔で前記監視を行う

ことを特徴とするプロセスの実行多重度制御処理方式。

3. 発明の詳細な説明

(概 要)

待ち要求の数等の変化に応じて実行多重度を変更するプロセスの実行多重度制御処理方式に関し、

プロセスについての待ち要求の数及び空きスレッドの数の変化に応じてプロセスの実行多重度を動的に変更可能と

することを目的とし、

プロセスについての待ち要求の数と空きスレ

ドの数をフィードバック機構部において求められた監視時間間隔で監視するモニタリング機構部と、前記監視の結果に基づいてスレッドの増減数と新たな監視時間間隔とを求めるフィードバック機構部と、前記スレッドの増減数に従って前記スレッドの数を変更することによって、実行多重度を変更する多重度変更部とを設け、前記待ち要求の数の増加又は減少若しくは空きスレッドの数に応じて、前記多重度変更部が前記スレッドの数を増加又は減少させることにより前記実行多重度を大きく又は小さくすると共に、前記モニタリング機構部がより短い又はより長い監視時間間隔で前記監視を行うように構成する。

〔産業上の利用分野〕

本発明は、プロセスの実行多重度制御処理方式に関し、更に詳しくは、待ち要求の数の変化に応じて実行多重度を調整するプロセスの実行多重度制御処理方式に関する。

データ処理システムの導入及び運用にあたって

- 3 -

てられる。一方、プロセスに対する要求の数が実行多重度の数即ちスレッドの数を越えると、当該上回った分の要求が処理待ちの状態とされる。

〔発明が解決しようとする課題〕

前述の従来技術によれば、プロセスの実行多重度を動的に変更することができない。このため、プロセスに対する負荷が実行多重度より大きい場合には、処理待ち要求が増加してシステム全体のスループットが低下してしまう。逆に、プロセスに対する負荷が実行多重度より小さい場合には、空きスレッド（要求を実行していないスレッド）が増加してシステム資源の使用効率が低下してしまう。

一方、このような問題があっても、実時間処理を必要とするアプリケーションの実行に際して、システムパラメタを変更して対応することは、システムファイルの書き換えを伴うため実質的に不可能である。

また、以上のような理由から、プロセスの実行

は、当該システムの処理のスループットを向上させ、かつこれを一定に保つことが必要となる。このスループットはプロセスの実行多重度をどのように設定するかによって左右される。

〔従来の技術〕

従来はプロセスがある一定の時間内で処理可能な最大の要求受付数を設定することにより、プロセスの実行多重度を設定していた。

この設定は、外部から人手により、前記最大の要求受付数をシステムパラメタとしてシステムファイルに設定することによって行われる。即ち、実行多重度は静的（固定的）に設定される。

プロセスは、その実行多重度の数だけの要求を、当該システムの資源を用いて、タイムシェアリングにより並列して（多重に）実行する。従って、プロセスは、その実行多重度の数だけの実行単位（以下スレッドという）を持ち、スレッドの数だけの要求を多重に実行すると言うことができる。このために、各スレッドにはシステム資源が割当

- 4 -

多重度の設定については、スループットやシステム資源の低下をできるだけ良くするために最大負荷予測値に基づいて調整することが必要となり、経験を積んだ人間にしかできないという問題を生じていた。そして、このことは、データ処理システムの導入から運用開始までに必要な作業の中でボトルネックとなっていた。

本発明は、プロセスについての待ち要求の数及び空きスレッドの数の変化に応じてプロセスの実行多重度を動的に変更可能としたプロセスの実行多重度制御処理方式を提供することを目的とする。

〔課題を解決するための手段〕

第1図は本発明の原理構成図であり、本発明によるデータ処理システムを示している。

第1図において、1はオペレーティングシステム（OS）、2はモニタリング機構部、3はフィードバック機構部、4は多重度変更部、5はプロセス、51は要求受付部、61ないし64はスレッド、7はプロセス5についての要求である。

オペレーティングシステム（又はそのプロセス管理部）1は、プロセス5の生成及び消滅、プロセス5におけるスレッド61ないし64の増減、各スレッド61ないし64へのシステム資源の割当て等を行う。

プロセス5は、要求（処理要求）7を実行する主体であり、要求7を多重に実行可能なものとされ、また、その実行多重度は自己についての待ち要求の数（の廃化）に応じて動的に変更可能とされる。即ち、プロセス5は、実行多重度の数だけのスレッド61ないし63を持ち、スレッド61ないし63の数だけの要求7を多重に実行する。

処理要求があった（要求7がプロセス5に到着した）時点での実行多重度を越える分の要求7は、例えば要求受付部51にキューイングされ、待ち要求（処理待ちの状態）とされる。一方、（実行可能な）要求7の数が実行多重度より小さい場合には、要求7を実行していない空きスレッドが生じる。空きスレッドを含めて、各スレッド61ないし63には、システム資源が割当てられる。

- 7 -

で、プロセス5の監視を行う。

〔作用〕

モニタリング機構部2による監視の結果、待ち要求の数が増加している時（又は多い時）には、フィードバック機構部3はスレッドの数を増し監視時間間隔を短くするようにこれらの値を求める。これにより、多重度変更部4がスレッドの数を増加させて実行多重度を大きくすると共に、モニタリング機構部2がより短い監視時間間隔でプロセス5の監視をする。従って、プロセス5の負荷が大きい（要求7の数が多）場合であっても、スループットの低下を防ぎ、システム全体のスループットを略一定に保つことができる。また、急激な負荷の増加にも十分対応できる。

逆に、待ち要求の数が減少している時（又は少ない時）若しくは空きスレッドが存在している時には、スレッドの数が減少させられることにより実行多重度が小さくされると共に、より長い監視時間間隔での監視が行われるようにされる。これ

オペレーティングシステム1には、モニタリング機構部2、フィードバック機構部3及び多重度変更部4が設けられる。

モニタリング機構部2は、監視対象であるプロセス5についてその待ち要求の数及び空きスレッドの数を、原則としてフィードバック機構部3において求められた監視時間間隔で監視し、その結果をフィードバック機構部3に通知する。

フィードバック機構部3は、前記監視結果に基づいて、プロセス5のスレッドの増減数と新たな監視時間間隔とを求め、前者を多重度変更部4に通知し、後者をモニタリング機構部2に通知する。

多重度変更部4は、スレッドの増減数に従ってプロセス5のスレッドの数を変更する。例えば、既存のスレッド61ないし63の他に新たなスレッド64を追加起動し、その数を「4」に増加させる。このスレッド数の変更により実行多重度が変更される。（「3」から「4」に増加させられる）。

モニタリング機構部2は、新たな監視時間間隔

- 8 -

により、被監視プロセス5の持つ不要なスレッド（又は空きスレッド）が解放されるので、これらに割当てられていたシステム資源も解放され有効利用が可能となる。また、負荷が小さい場合における必要以上のモニタリング機構部2の動作を抑え、オペレーティングシステム1のオーバーヘッドを削減できる。

〔実施例〕

第1図図示のデータ処理システムについて、更に説明する。

モニタリング機構部2は、オペレーティングシステム1内部の常駐機能として存在し、所定の監視時間間隔で動作して被監視プロセス5を監視する。この監視の結果、プロセス5の待ち要求数及び空きスレッド数の情報を採取し、当該監視時間間隔と共にフィードバック機構部3へ渡す。

モニタリング機構部2は、被監視プロセス5の起動と共に動作を開始し、まず最初は当該システムでシステムパラメタとしてシステムファイルに

- 9 -

- 275 -

- 10 -

BEST AVAILABLE COPY

procedure フードバック機構

以下、フィードバック機構部3の行う処理について示す。

- 11 -

① まず、前処理として、「今回の持ち要求数の増減数」及び「今回の持ち要求増加率」を求める。このために、フィードバック編組部3は、「前回の持ち要求数」を保持する。

- 13 -

- 1 2 -

(四)「今回の持ち要求数の増減数」 > 0 、即ち、

- 14 -

待ち要求数が減少した場合；

(ア) 前述の(1)(ア)と同様に、待ち要求増加率が減少、変化なし、及び、増加の場合、各々、監視時間間隔を短く、変更なし、及び長くするように新しい監視時間間隔を求める。

(イ) 前述の(1)(イ)と同様に、スレッド数が待ち要求数の減少数だけ減少される。

③ 多重度変更部4を呼び出し、これにスレッド増減数を渡す。また、新しい監視時間間隔をモニタリング機構部2に渡す。

なお、以上の処理は待ち要求数に基づいたものであるが、これとは別に空きスレッド数に基づいた処理も行われる。例えば、空きスレッドが存在する場合、

スレッド増減数＝空きスレッド数

として、スレッド数を減少させる処理を行う。

また、以上のフィードバック機構部3における処理は、スレッド増加閾値を越える値がプロセス

- 15 -

荷が増加したと判断し、多重度変更部4に対してスレッド数を増加させることを、モニタリング機構部2に対して監視時間間隔を短くすることを依頼する。これにより、時刻 t_1 において、スレッド数が増加される。

時刻 t_1 までは、スレッド数を増加してもなお待ち要求数が増加しているため、スレッド数が重ねて増加される。

時刻 t_1 において、フィードバック機構部3は、待ち要求数が時刻 t_1 のときの値より減少していることを知る。これにより、多重度変更部4はスレッド数を減少させ、モニタリング機構部2は監視時間間隔を長くする。

なお、第3図において監視時間間隔は等しい間隔で示してあるが、実際は、時刻 t_1 までは監視時間変化レイトに従って順に短くされ、一方、時刻 t_1 以後は同様のレイトで順に長くされている。

また、第3図において、実線の折れ線は待ち要求数の変化を示し、点線は見掛け上の待ち要求数の変化を示す。

5の監視結果として得られた場合にのみ、行われるようにすることができる。

以上の処理による待ち要求数、空きスレッド数及びスループットの間の関係は、第2図に示す如くなる。即ち、待ち要求数の増加は空きスレッド数の減少を招き、空きスレッド数の増加はスループットの増加を招き、スループットの増加は待ち要求数の減少を招く。従って、この因果関係は負のフィードバックループを構成する。

第3図は、第1図図示のデータ処理システムにおけるスレッド数の変化、即ち、実行多重度の制御状態を示す図である。

ある時刻 t_1 において、待ち要求数は「0」で、スレッド数は図示の如く所定の(小さい)値とされている。

時刻 t_1 から所定の監視時間間隔だけ経過した時刻 t_2 において、モニタリング機構部2からの通知によりフィードバック機構部3は、待ち要求の数が時刻 t_1 のときの値より増加していることを知る。そこで、フィードバック機構部3は、負

- 16 -

以上の如く、スレッド数は各監視時刻において、待ち要求数の変化に応じて動点に変更される。

(発明の効果)

以上説明したように、本発明によれば、プロセスの実行多重度制御処理において、待ち要求数及び空きスレッド数の変化に応じて動的にスレッド数及び監視時間間隔を変更することにより、負荷の急激な変動に対応してシステム全体のスループットを向上しかつ一定に保つことができ、また、システム資源を有効に活用することができる。更に、以上の如き処理を自動的にシステム自体によって行うことができるので、システム導入から運用までの期間を短縮し、必要な工数を削減することができる。

4. 図面の簡単な説明

第1図は本発明の原理構成図、

第2図は因果関係を示す図、

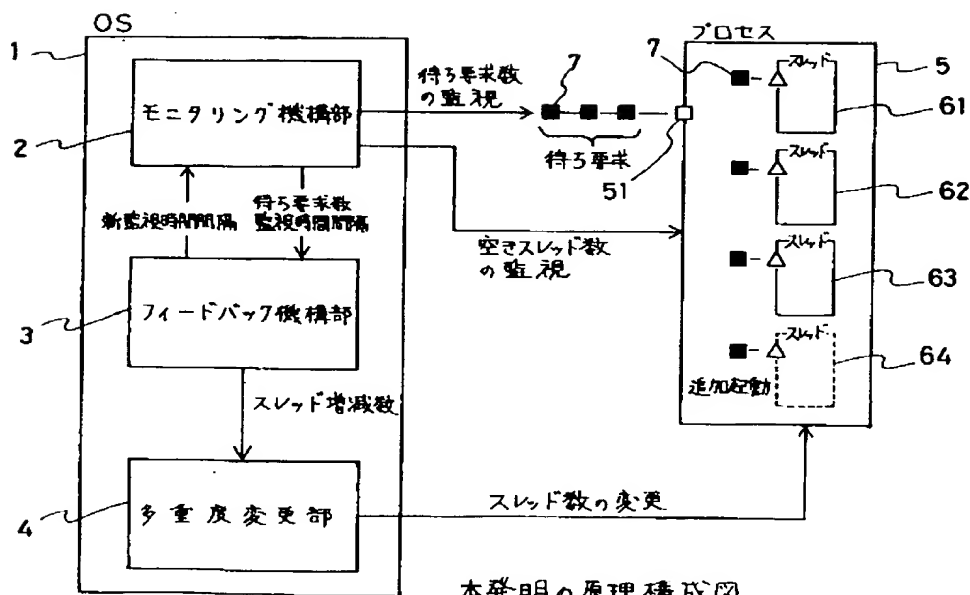
第3図は実行多重度の制御状態を示す図。

1 はオペレーティングシステム (OS)、2 は
モニタリング機構部、3 はフィードバック機構部、
4 は多重度変更部、5 はプロセス、51 は要求受
付部、61 ないし 64 はスレッド、7 はプロセス
5 についての要求である。

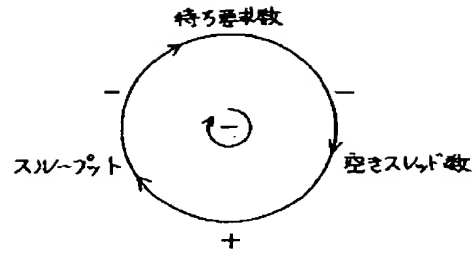
特許出願人 富士通株式会社

代理人 弁理士 森田 寛 (外 2 名)

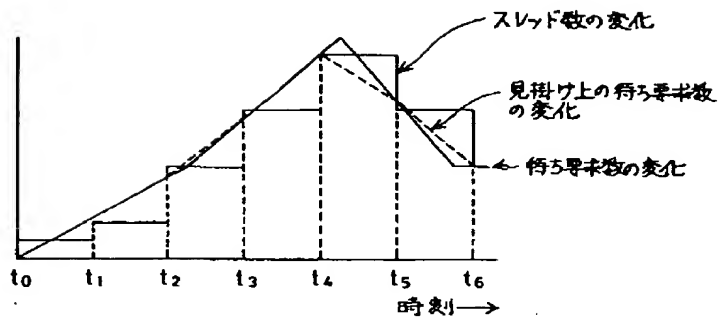
- 19 -



本発明の原理構成図
第 1 図



因果関係を示す図
第 2 図



実行多重度の制御状態を示す図
第 3 図